Cloning of Tumor-associated Differentially Expressed Gene-14, a Novel Serine Protease Overexpressed by Ovarian Carcinoma


Departments of Obstetrics and Gynecology [T. O., T. H. P.] and Biochemistry and Molecular Biology [L. J. U., Y. W., T. J. O.], University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205-7199; and Department of Obstetrics and Gynecology, Hiroshima School of Medicine, Hiroshima 734, Japan [K. S.]

ABSTRACT

The family of enzymes known as serine proteases supports many biological functions for cancer cells, including activation of growth and angiogenic factors and activation of other proteases for invasion and metastasis. In addition, many of these serine proteases are secreted by cells into the extracellular space to serve these functions. Therefore, serine proteases are excellent candidate tumor markers. To examine serine proteases expressed by ovarian carcinoma, we designed degenerate PCR primers corresponding to the conserved regions of these genes and used them in reverse transcriptase-PCR experiments with normal and tumor cDNA as a template. The PCR products were subcloned and sequenced, and one of these clones was found to encode a novel serine protease, named tumor-associated differentially expressed gene-14 (TADG14). Northern blot analysis indicated that the mRNA for TADG14 is 1.4 kb long and that it is highly overexpressed in ovarian carcinoma compared with normal ovary. The entire cDNA has been obtained, and based on sequence homology, it encodes a 260-amino acid serine protease. Semi-quantitative PCR indicates that TADG14 is overexpressed in 24 of 40 tumors studied. Northern blot data confirm this overexpression, and immunohistochemical staining suggests that this protein is secreted. As such, the TADG14 protease may be useful as a diagnostic tool or as a molecular target for therapy.

INTRODUCTION

Serine proteases comprise a family of protein-degrading enzymes that serve as a host of biological functions, including activation of blood coagulation cascades, activation of growth and angiogenic factors, and degradation of extracellular matrix components (1–4). In recent years, aberrant expression of serine proteases such as plasminogen activator has been shown to correlate positively with the invasiveness and metastatic potential of tumor cells (3, 5, 6). Presumably, this occurs because the ability of the tumors to degrade extracellular matrix components is increased, either directly or indirectly through the proteolytic activation of other zymogenic proteases. More significantly, the serine protease known as PSA has been used successfully as a tumor marker for the early diagnosis of prostate cancer due to its abnormal prevalence in the peripheral blood of these patients (7). Serine proteases play important roles in the cascade of events involved in the malignant process, and at least for prostate cancer, they provide sufficient signal to allow detection of early disease.

This year, the American Cancer Society predicts that there will be 14,500 ovarian cancer-related deaths in the United States and that there will be 25,400 new cases diagnosed (8). Unfortunately, although diagnostic assays based on the ovarian cancer antigen CA125 have improved physicians’ ability to diagnose and monitor recurrence of ovarian carcinoma, the problem remains that most of these new cases will be diagnosed at late stages, in which the primary tumor has progressed to a metastatic state. This failure to diagnose patients in the early stages of ovarian cancer directly impacts the 5-year survival rate for ovarian cancer patients, which remains below 50% (8). This study is the result of a strategy to identify those serine proteases that are overexpressed by ovarian carcinomas in an effort to define potential tumor markers.

All serine proteases contain conserved histidine, aspartate, and serine residues that are necessary for enzymatic activity. To identify the expressed serine proteases, we used degenerate oligodeoxynucleotide primers designed to the conserved amino acid sequences surrounding the invariant His and Ser residues of the catalytic triad (9) in PCRs with cDNA from either normal ovarian tissue or ovarian carcinoma as the template. PCR products of the appropriate size were subcloned into T-vector and sequenced. Previously, this strategy has proved successful in identifying the serine proteases hepsin and straturn corneum chymotryptic enzyme, which have been shown to be expressed at abnormally high levels in ovarian carcinoma (10). Homology searches revealed that one of the subclones obtained from ovarian carcinoma represented a novel 406-bp sequence that has significant sequence similarity to other known proteases, including mouse neuropsin, hHk2, and human PSA. The complete cDNA for this novel sequence was cloned and found to encode a trypsin like serine protease, named TADG14. More importantly, the TADG14 transcript was found to be highly expressed in a majority of ovarian tumors but not expressed by normal ovarian tissue. High-level expression of TADG14 appears to be restricted to tumors, and this protease appears to be secreted in a manner that would suggest a possible role in invasion and metastasis. Moreover, due to the extra-cellular nature of this enzyme, it may be possible to exploit its expression as a diagnostic tool for ovarian cancer.

MATERIALS AND METHODS

**mRNA Isolation and cDNA Synthesis.** Forty-one ovarian tumors and 10 normal ovary specimens were obtained from surgery and frozen in liquid nitrogen or purchased from the Cooperative Human Tissue Network. mRNA isolation was performed according to the manufacturer’s instructions using the oligo(dT) chromatography-based Mini RiboSep Ultra mRNA isolation kit (Becton Dickinson, Bedford, MA).

First strand cDNA was synthesized using 5.0 μg of mRNA and either random hexamer or oligo(dT) primers according to the manufacturer’s protocol using a first strand synthesis kit (Clontech, Palo Alto, CA). The purity of the cDNA was evaluated by PCR using primers specific for the p53 gene. These primers span an intron such that the PCR products generated from pure cDNA can be distinguished from cDNA that is contaminated with genomic DNA.

**PCRs.** Reactions with degenerate primers and quantitative PCRs were carried out as described previously (10). The sequences of the TADG14-
Serine Protease and Ovarian Cancer

Fig. 1. Northern blot analysis. A, mRNA was isolated from the tissues of interest as described (10) and subjected to Northern hybridization using a random-labeled 230-bp TADG14-specific RT-PCR product. The blot was stripped and probed for β-tubulin. B-D, MTN blots (Clontech) were probed with the same TADG14- and β-tubulin-specific RT-PCR products. TADG14 mRNA was detected as a 1.4-kb transcript in tumors and was not detected in any normal tissue studied.

Western blot analysis, (T14-1), GHECQPHSQPWQ (T14-2), and LDWIKKIIGSKG (T14-3). For Western blot analysis, ~20 μg of MDA-MB-435 phosphatase-depleted HeLa cell lysates were size-separated by electrophoresis through a 6.3% formaldehyde-1.2% agarose gel in 0.02 M MOPS, 0.05 M sodium acetate (pH 7.0), and 0.001 M EDTA. The mRNAs were then blotted to Hybond-N (Amersham, Piscataway, NJ) by capillary action in 20× SSPE. The mRNAs were fixed to the membrane by baking for 2 hours at 80°C. Additional MTN blots were probed with Clontech. These blots include the human MTN blot, the human fetal MTN II blot, and the human brain MTN III blot.

DNA Sequencing. Using a plasmid-specific primer near the cloning site, we carried out sequence reactions using Prism Ready Reaction Dye Deoxy terminators (Applied Biosystems, Foster City, CA) according to the manufacturer’s instructions. Residual dye terminators were removed from the completed sequencing reaction using a Centri-sep spin column (Princeton Separation, Adelphia, NJ). An Applied Biosystems model 373A DNA Sequencing System was used for sequence analysis. Sequences were compared with the Genetics Computer Group Pileup program, the amino acid sequence of TADG14 was compared with hHk2 (accession no. P06870), human PSA (hPSA, accession no. P05278), mouse neuropsin (mNeur, accession no. P07288), and human protease M (hProM, accession no. U62801). Black shading, amino acid residues that are identical in at least three of the five sequences. Gray shading, amino acid residues that are similar among at least three sequences. 

ANTIBODY PRODUCTION AND WESTERN BLOT ANALYSIS. Polyclonal antibodies were generated by immunization of white New Zealand rabbits with one of three poly-lysine-linked multiple antigen peptides derived from the deduced amino acid sequence of TADG14. These sequences are KYTVRLGDHSLQ (A), THTRGSQGLQSG (B), and CVTVIYSLDV (C). For Western blot analysis, ~20 μg of MDA-MB-435 phosphatase-depleted HeLa cell lysates were size-separated on a 15% SDS-polyacrylamide gel and electroblotted to polyvinylidene difluoride at 100 V for 40 min at 4°C. The blot was blocked overnight in Tris-buffered saline (pH 7.8) containing 0.2% nonfat milk. Primary antibody was added to the membrane at a dilution of 1:100 in 0.2% milk-Tris-buffered saline and incubated for 2 hours at room temperature. The blot was washed and incubated with a 1:3000 dilution of alkaline phosphatase-conjugated goat antirabbit IgG antibody (Bio-Rad, Hercules, CA) for 1 hour at room temperature. The blot was washed and incubated with a chemiluminescent substrate (Bio-Rad) before a 10-s exposure to X-ray film for visualization.
buffer (pH 6.0). The specimens were incubated in methanol with 0.3% H$_2$O$_2$ and stained with ethidium bromide. In this figure, the 454-bp band represents the experiment. The reaction products were electrophoresed through a 2% agarose gel and products were quantitated as described (10).

Fig. 3. TADG14 quantitative PCR. A, typical results of a TADG14 quantitative PCR experiment. The reaction products were electrophoresed through a 2% agarose gel and stained with ethidium bromide. In this figure, the 454-bp band represents the β-tubulin product, and the 230-bp band represents the TADG14 product. The radiolabeled PCR products were quantitated as described (10).

B

Fig. 4. Western blot. Polyclonal antibodies were generated by immunization of rabbits with one of three poly-lysine-linked multiple antigen peptides derived from the deduced amino acid sequence of TADG14. For Western blot analysis, ~20 μg of MDA-MB-435S and HeLa cell lysates were separated on a 15% SDS-polyacrylamide gel and electroblotted to polyvinylidene difluoride. After probing with either of two of the peptide-derived antibodies, a single M, 30,000 band was observed. This is in close agreement with the expected size of the TADG14 protein.

**RESULTS**

After confirming that the 406-bp PCR product was unique and was appropriately conserved to fit into the serine protease family, we used this PCR product as a probe for Northern blot analysis to determine the transcript size and tissue specificity of its expression. It was found that the mRNA for this clone is ~1.4 kb (Fig. 1A) and that it is strongly expressed in ovarian carcinomas but not in normal ovary. More importantly, the transcript was found to be undetectable by Northern analysis in 28 normal human tissues studied (Fig. 1, B–D; data not shown). In a more sensitive assay of 50 normal human tissues (Clontech), RNA dot blot analysis revealed that this clone was very weakly expressed in only 3 of these 50 tissues: the kidney, the lung, and the mammary gland (data not shown).

Using standard hybridization techniques, we screened a cDNA library that was constructed from the mRNA isolated from the ascites cells of an ovarian cystadenocarcinoma patient. Five clones were obtained, two of which overlapped and spanned 1343 nucleotides (Fig. 2A). The last two nucleotides prior to the poly(A) tail and the poly(A) tail itself were obtained from the National Center for Biotechnology Information EST database (accession no. AA343629).

Subsequent Northern blot analyses with probes derived from sequences near the 5’ or 3’ end of this cDNA were consistent with previous results suggesting that the obtained clones were produced by the same gene (data not shown). This cDNA includes a Kozak’s consensus sequence for the initiation of translation and a polyadenylation signal. The mRNA provides an open reading frame of 260 amino acids, which contains the necessary residues (His$^3$, Asp$^{120}$, and Ser$^{212}$) in the appropriate context to classify this protein as a trypsin-like serine protease (11). Near its NH$_2$ terminus, the predicted protein contains a stretch of hydrophobic amino acids that probably serve as a secretion signal sequence (12). In addition, residues 110–112 encode a potential site for glycosylation that is common to serine proteases of the kallikrein subfamily, such as PSA. This enzyme was named TADG14, and the sequence was submitted to GenBank (accession no. AF055982).

Comparison of the deduced TADG14 amino acid sequence with sequences of known proteases revealed that it possesses significant similarity with hHK2, PSA, protease M, and mouse neuropsin (13–16). At the amino acid level, TADG14 is 48% identical to protease M.

**Table 1** TADG14 overexpression by tissue subtype

<table>
<thead>
<tr>
<th>Tissue type</th>
<th>TADG14 overexpression$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0/10 (0%)</td>
</tr>
<tr>
<td>LMP</td>
<td>4/10 (40%)</td>
</tr>
<tr>
<td>Serous</td>
<td>3/6 (50%)</td>
</tr>
<tr>
<td>Mucinous</td>
<td>1/4 (25%)</td>
</tr>
<tr>
<td>Carcinoma</td>
<td>20/30 (67%)</td>
</tr>
<tr>
<td>Serous</td>
<td>13/17 (76%)</td>
</tr>
<tr>
<td>Mucinous</td>
<td>1/4 (14%)</td>
</tr>
<tr>
<td>Endometrioid</td>
<td>3/3 (100%)</td>
</tr>
<tr>
<td>Clear cell</td>
<td>3/3 (100%)</td>
</tr>
</tbody>
</table>

$^a$ Overexpression was defined as ≥2SD over the mean normal values.
46% identical to hHk2, and 43% identical to PSA. More interestingly, the mouse protease neuropsin and TADG14 share 72% amino acid identity (Fig. 2B). In addition to the similarity of the protein sequences, neuropsin and TADG14 mRNAs are of similar size (1.4 kb) and structure, with approximately the same amounts of 5’ and 3’ untranslated regions, suggesting the possibility of orthology. Neuropsin was originally identified as being expressed in mouse hippocampus and shown to be differentially expressed under stimulation (16). However, TADG14 mRNA was undetectable in human whole brain by Northern blot. Furthermore, Northern blot analysis for TADG14 in eight separate parts of human brain, including amygdala, caudate nucleus, corpus callosum, hippocampus, whole brain, substantia nigra, subthalamic nucleus, and thalamus, also turned out to be negative. Recently, a human cDNA encoding neuropsin was submitted to the GenBank database (accession no. AB009849). Although this clone represents a different transcript from TADG14, it encodes a protein that is identical to TADG14 (17). Therefore, it seems logical that TADG14 and neuropsin may arise as alternative splicing products from the same gene.

To characterize the extent and frequency of expression of the TADG14 gene in ovarian tumors, we used semiquantitative PCR with cDNA derived from normal ovary, ovarian carcinoma, or LMP tumors as template. This technique has been previously authenticated and verified by Northern blot, Western blot, and immunohistochemistry (10, 18). PCR primers that amplify a TADG14-specific 230-bp product were synthesized and used simultaneously in reactions with primers that produce a specific 454-bp PCR product for β-tubulin. A radiolabeled nucleotide was included in this reaction, the PCR products were separated on a 2% agarose gel, and the intensity of each band was quantitated by a PhosphoImager (Molecular Dynamics). Fig. 3A shows an ethidium bromide-stained agarose gel with the separated quantitative PCR products and is representative of the typical results observed.

The ratio of the TADG14 PCR product to that of β-tubulin (mean ± SD) was calculated for normal ovary (0.034 ± 0.024) samples, which all showed relatively low expression levels. TADG14 overexpression was defined as exceeding the mean of the ratio of TADG14 to β-tubulin for normal samples by >2 SDs. TADG14 was found to be overexpressed in 4 of 10 LMP tumors (40%) and 20 of 30 ovarian carcinomas (67%) studied. For individual histological subtypes of tumor, the expression ratio was 0.110 ± 0.092 for serous LMP tumors, 0.096 ± 0.142 for mucinous LMP tumors, 0.457 ± 0.345 for serous carcinomas, 0.171 ± 0.300 for mucinous carcinomas, 0.308 ± 0.144 for clear cell carcinomas, and 0.485 ± 0.325 for endometrioid carcinomas. Of the 30 ovarian carcinomas studied, 13 of 17 serous tumors, 1 of 7 mucinous tumors, 3 of 3 clear cell tumors, and 3 of 3 endometrioid tumors overexpressed TADG14 (Fig. 3B). These data are summarized in Table 1. Although not quantitated, transcripts for TADG14 were detectable in prostate and colon carcinoma (data not shown).

Immunogenic poly-lysine-linked multiple-antigen peptides were synthesized based on the deduced amino acid sequence of TADG14 and used to immunize rabbits for the production of polyclonal antibodies. The antisera raised to the peptide sequence LDWIKKIIG-SKG near the COOH terminal (amino acids 249–260) was used in Western blot analysis to determine whether this antibody would recognize a protein of the predicted size of M, 28,000. Proteins from the HeLa cell line and the carcinoma-derived MD-MB435S cell line were used in this experiment, and it was found that the antibody recognized a single M, 30,000 protein in both (Fig. 4, two right lanes). This size is within a reasonable range of the predicted molecular weight. As a negative control, duplicate HeLa and MD-MB435S lysates were examined with rabbit preimmune serum (Fig. 4, left two lanes). More importantly, this experiment was reproducible with antisera to a peptide from a different region of TADG14, suggesting that these cultured cancer cells produce the TADG14 protein.

Immunohistochemical staining supported the data obtained by quantitative PCR and by Northern blot. Using a TADG14 peptide-directed antibody (T14-1; described above), we observed no staining with normal ovarian tissue samples (Fig. 5A). However, intense staining was associated with tumor cells of all of the various histological subtypes of ovarian carcinoma examined. For serous carcinoma (Fig. 5, B and C), the antigen appears to be associated with tumor cells in the form of granules. These granular structures may be intermediates in the pathway that ultimately leads to secretion of TADG14. In mucinous and clear cell carcinoma samples (Fig. 5, D and F, respectively), TADG14 is highly associated with the tumor cells. In endometrioid carcinoma (Fig. 5E), the antigen is most prevalent in the glandular lumen formed by the tumor cells.

DISCUSSION

The lethality of neoplastic cells lies in their ability to proliferate abnormally and invade normal host tissues. Malignancies use proteases to provide a variety of services that assist in the process of tumor progression, including activation of growth and angiogenic factors, and to provide the basis for invasion and metastasis. In the process of studying these enzymes, we have identified overexpression of the known proteases, hepsin and stratum corneum chymotryptic enzyme. In this study, we have cloned a cDNA encoding a novel serine protease, TADG14. This protease was found to be very highly expressed in 67% (20 of 30) of ovarian carcinomas studied, whereas
it was undetected in normal ovarian tissue. We were also unable to detect the TADG14 transcript in any of 50 normal human tissues studied. Upon prolonged Northern blot exposure, extremely low levels of TADG14 were detected in normal kidney, breast, and lung. This suggests the possibility that this gene is under the control of a promoter that is most active in ovarian tumors, and it may be possible to exploit this for therapeutic means. Unfortunately, TADG14 expression can be detected in other types of cancer, including prostate, breast, and colon. This may limit the usefulness of TADG14 as a potential diagnostic marker for ovarian carcinoma, but it in no way detracts from the usefulness of this molecule as a target for cancer therapy or the usefulness of the TADG14 promoter in gene therapy applications.

At the nucleotide level, TADG14 mRNA resembles the recently cloned human neurosin transcript, with obvious differences residing in the 5′ and 3′ UTRs. TADG14 mRNA contains 491 bases of 5′ UTR that were not found in human neurosin. Also, the nucleotides preceding the poly(A) tail in the 3′ UTR are not homologous. A 0.9-kb transcript for human neurosin was identified in cultured keratinocytes but not in normal hippocampus. Also, it was not identified as being associated with tumors. At the amino acid level, TADG14 is identical to human neurosin.

Among other known proteases, TADG14 most closely resembles the mouse protease known as neurosin, which was originally cloned from mouse hippocampus and, subsequently, implicated in neuronal plasticity (19). If TADG14 functions in a manner similar to mouse neurosin, it may be capable of restructuring the three-dimensional architecture of a tumor, allowing for shedding of tumor cells or invasion of normal host tissues by degrading fibronectin (20). In support of this, immunohistochemical staining of ovarian tumors shows TADG14 is highly associated with tumor cells and the cells near the invasive fronts of tumor. Therefore, TADG14 could be an important target for the inhibition of tumor progression.

Most importantly, the 5-year survival rate for ovarian cancer patients remains <50% because of an inability to diagnose this disease at an early stage. TADG14 contains a secretion signal sequence and immunohistochemical data suggest that TADG14 is secreted. In addition, by Northern blot and RNA dot blot analyses, TADG14 appears only in abundance in tumor tissues. As a result of this, it may be possible to design assays based on the detection of this protein for the early detection of ovarian cancer. Currently, the best available ovarian cancer tumor marker is CA125. However, due to high endogenous circulating levels of this antigen, the signal:noise ratio limits its usefulness as a diagnostic tool. Therefore, TADG14, due to its limited expression in other tissues and potential for being present in the circulation of tumor-bearing patients, may prove to be a useful tool for early detection of ovarian cancer, especially the most prevalent serous cystadenocarcinoma subtype.

ACKNOWLEDGMENTS

We acknowledge the participation of the following divisions of the Cooperative Human Tissue Network in providing tumor tissues: Western Division, Case Western Reserve University (Cleveland, OH); Midwestern Division, Ohio State University (Columbus, OH); Eastern Division, National Disease Research Interchange (Philadelphia, PA); Pediatric Division, Children’s Hospital (Columbus, OH); and Southern Division, University of Alabama at Birmingham (Birmingham, AL).

We also acknowledge Yan Yan for technical assistance in the original subcloning experiments.

REFERENCES

Cloning of Tumor-associated Differentially Expressed Gene-14, a Novel Serine Protease Overexpressed by Ovarian Carcinoma
